

2021

Usability of Sentinel-3 satellite ocean colour data for examination and chlorophyll gradient classification of submesoscale island wake formations

Farrington, Emmanuelle Simone

Farrington, E.S. (2021) 'Usability of Sentinel-3 satellite ocean colour data for examination and chlorophyll gradient classification of submesoscale island wake formations', *The Plymouth Student Scientist*, 14(1), pp. 49-77.

<http://hdl.handle.net/10026.1/17337>

The Plymouth Student Scientist
University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Usability of Sentinel-3 satellite ocean colour data for examination and chlorophyll gradient classification of submesoscale island wake formations

Emmanuelle Simone Farrington

Supervisor: Dr Jill Schwarz, School of Biological and Marine Sciences

Room 111 Reynold, Drake Circus, Plymouth, PL4 8AA

Abstract

Submesoscale turbulent formations present challenges to sampling efforts due to their scale and transient nature. The development of successful sampling methods is imperative to achieving an understanding of the biological and ecological implications of these turbulent formations. This study assesses the viability of using Sentinel-3 OLCI multispectral satellite data for the study of submesoscale island wake formations. This research subsequently samples the eddies and filaments with virtual transects of satellite chlorophyll products, calculates metrics such as formation gradient and FWHM, identifies key environmental drivers of wake characteristics, and presents a novel framework by which to classify these formations. The study is subject to limitations regarding cloud cover of a small area of interest, but nonetheless is able to identify seasonal occurrence of formations, and delineate environmental variables driving formation metrics (i.e. eddy gradient is significantly correlated with current velocity (MLR; slope = 1.490, p-value = 0.001) and current angle (MLR; slope = 1.660×10^{-02} , p-value = 0.016), and eddy FWHM is significantly correlated with: a formation's distance from island (MLR; slope = 0.015, p-value = 0.024), island diameter (MLR; slope = 0.361, p-value = < 0.001), and current velocity (MLR; slope = 10.205, p-value = < 0.001)). An eddy and filament characterisation framework is proposed regarding chlorophyll gradient, in anticipation of the biological implications of such gradients: weak gradient (< 25th percentile of all gradient values of the formations sampled within the AOI), moderate gradient (> 25th percentile and < 75th percentile of all gradient values), and strong gradient (> 75th percentile of all gradient values) formations. The results regarding environmental drivers of wake characteristics align with previous studies of the physical implications of island wakes, as well as provide new insight. This research presents a missing link between a physical understanding of submesoscale island wakes and a more extensive understanding regarding the biological signatures – expanding the potential for further study regarding the biological and conservation/management implications.

Keywords: Kuroshio Island Wakes, Submesoscale Circulation, Ocean Colour, Satellite Remote Sensing, Chlorophyll

Introduction

Phytoplankton are the primary producers of the ocean. Although specific growth-limiting nutrients and minerals may differ between functional phytoplankton groups and geographic regions, two bloom-inducing drivers are universal: light and nutrients. The concurrence of light and nutrients arises through a variety of natural processes. In the coastal ocean, nutrients are made available by terrigenous runoff, erosion, and coastal upwelling. In the open oligotrophic ocean, nutrient supply to the photic zone relies on oceanographic processes to raise the deep nitracline. These processes include wind-driven mixing of stratified surface waters, internal waves, tidal fronts, mesoscale eddy upwelling, and island dynamics resulting in island wakes or eddy trains.

The present research investigates surface signatures of island dynamic driven phytoplankton production. Associated phenomena can occur on a range of scales. Islands have been shown to generate diffuse chlorophyll wakes covering 25,000 km² (Palacios 2002), concentrated wake trails on a scale of several kilometres (Suthers *et al.* 2004), seasonal or semi-permanent mesoscale eddy features (Cherubin and Garavelli 2016), and Karman vortex-like submesoscale eddy trains (Isoguchi *et al.* 2009). Submesoscale eddies have only recently risen in prominence within remote sensing literature; previous schools of thought recognised the importance of mesoscale oceanographic processes, but submesoscale features were widely left unstudied.

Unlike mesoscale eddies, submesoscale eddies are not geostrophically balanced (Hosegood *et al.* 2017; Mahadevan and Tandon 2006; Zhang *et al.* 2019). This means they are defined by having a diameter less than the Rossby radius of deformation (Zhang 2019), wherein buoyancy effects exceed the effects of rotational motion (geostrophy). Submesoscale eddies form when wind stress or a converging flow field induces frontogenesis, disrupting a geostrophically balanced surface ocean by producing steep horizontal pressure gradients (Hosegood *et al.* 2017; Mahadevan and Tandon 2006; Zhang *et al.* 2019). Horizontal shear is strong enough to induce turbulent energy dissipation in the form of eddies and vortices where Richardson number (Ri) approaches 1 (Hosegood *et al.* 2017; Mahadevan and Tandon 2006; Zhang *et al.* 2019). Submesoscale eddies have evolution timescales generally within 1 day, and generate stronger vertical velocities than mesoscale eddies (Hosegood *et al.* 2017; Mahadevan and Tandon 2006), reaching depths of up to 1 km. These strong vertical currents drive surface phytoplankton production (Mahadevan 2015), due to an enhanced ability to break the deep pycnocline and draw up nutrients.

The methodology for sampling submesoscale eddies remains uncertain, and the techniques used for mesoscale sampling may not be viable at submesoscale. Chen (2004) used a Lagrangian particle transport model to identify eddies along the Kuroshio current. The model successfully identified mesoscale frontal eddies where the Kuroshio enters the North Pacific, but it failed to resolve any submesoscale features (Chen 2004). These data could be supplemented using *in situ* drifter buoys with fluorometers in an effort to resolve submesoscale dynamics and chlorophyll signatures, however this method is costly. Godo *et al.* (2012) collected conductivity, temperature and depth (CTD) profiles, acoustic doppler current profiler (ADCP), and echosounder measurements using a star-shaped transect pattern across a predictable mesoscale eddy formation. The echosounder revealed fish aggregations feeding along the density gradients of the eddy (Godo *et al.* 2012), immediately confirming biological

importance of the eddy – something that remains uncertain for submesoscale equivalents. Although this sampling method was successful for a predictably occurring mesoscale eddy, it may not be suitable for confirming biological relevance of a highly transient and unpredictable submesoscale equivalent. Attempting star-shaped transects of a submesoscale eddy would likely be difficult to perform successfully in the face of high current velocities and downstream propagation.

Remote sensing may provide a solution to the submesoscale biological sampling problem. The development of remote sensing via satellite introduced a whole host of indicators (objective metrics) with which to monitor the marine environment (Platt *et al.* 2009). Because phytoplankton affect the optical characteristics of the ocean surface through photosynthetic pigment absorption and scattering, multispectral sensors coupled with well-calibrated ocean colour algorithms can be used to estimate surface chlorophyll (CHL) concentrations as a proxy for phytoplankton biomass. Ocean colour sensors have been in orbit since the launch of the Coastal Zone Colour Sensor (CZCS) in 1978 (Groom *et al.* 2019). Subsequent satellites providing global coverage from ocean colour sensors include: NASA's Seastar-SeaWiFS with up to 1100 m resolution and 8 spectral bands, ESA's Envisat-MERIS with 1000 m resolution over ocean and 300 m over land and coast with 15 spectral bands, NASA's Terra/Aqua-MODIS with 1000 m resolution in ocean colour bands and 9 bands in the visible/NIR (Groom *et al.* 2019; Morel *et al.* 2007), and NASA's LANDSAT missions with 300 m resolution (Isoguchi *et al.* 2009). Multispectral data from these satellites, processed with 3-band and 4-band ratio algorithms (Groom *et al.* 2019), have been used to identify numerous oceanographic features. However, because submesoscale eddies have diameters of 1-10 km, historic ocean colour data has not been of sufficient spatial resolution to resolve these particular features. Recent developments in satellite ocean colour missions include ESA's Sentinel-3 A and B satellites, launched in 2016 and 2018 respectively, both of which carry the Ocean Land and Colour Instrument (OLCI). Although LANDSAT achieved 300 m spatial resolution, it lacked repeat imaging on a finer temporal scale than 12 days (Isoguchi *et al.* 2009). Therefore, Sentinel-3 data provides the highest coupled spectral and spatial resolution yet available for ocean colour data, with 300 m global resolution in 21 spectral bands, and with repeat imaging 2 times per day. The Sentinel-3 OLCI platform may provide the best possible data available for studying submesoscale chlorophyll island wake features.

Alongside these new developments, it is important to keep in mind that ocean colour products are only as good as the algorithms that produce them from raw multispectral imagery, and the data used for algorithm validation. Throughout the lifetime of satellite remote multispectral sensors, associated ocean colour algorithms have been developed. Band-ratio algorithms use the ratio of two or more spectral bands to estimate chlorophyll and are constructed by regression of remotely sensed data against *in situ* data (Groom *et al.* 2019). These generally use the blue-green spectral bands because phytoplankton absorption is primarily within these wavelengths. This type of algorithm is suitable for open ocean (Case 1) waters, containing only a single type of optically active constituents (chlorophyll; Antoine 2010). For complex (Case 2) turbid coastal waters, where signature interference by total suspended matter (TSM) or coloured dissolved organic matter (CDOM) is likely to occur due to fluvial input or resuspended particles (Blondeau-Patissier *et al.* 2014; Groom *et al.* 2019), bio-optical models are generally more suitable (Doerffer 2010). These models are based thoroughly on marine optical theory and are less likely to misconstrue CDOM and TSM signals as those of chlorophyll (Blondeau-Patissier *et al.* 2014).

Sentinel-3 Level 2 chlorophyll data products are available based on processing by two different algorithms. The OC4Me band ratio semi analytical algorithm, developed from the MODIS OC3M algorithm (O'Reilly *et al.* 1998), uses ratios of the 560, 510 and 443 nm bands (Antoine 2010). The Neural Network (NN) algorithm is a bio-optical model using 15 spectral bands (Doerffer 2010). OC4Me is generally restricted to Case 1 waters, and NN is suitable for turbid coastal waters (Antoine 2010; Doerffer 2010). The type of algorithm used for study will depend on which is better suited for a given region of interest. Because both are globally calibrated, it is possible that data for some sub-regions may not be accurate. It is prudent to validate algorithms for a specific area of interest in order to establish regional and local usability, because algorithm suitability varies between locations (Gregg and Casey 2004).

The Tokara islands, which lie in the Ryuku island chain at the boundary of the East China Sea and the Northwest Pacific, present an ideal case study for the examination of submesoscale features using Sentinel-3 data. The Tokara island chain transects the main axis of the Kuroshio, the western boundary current of the North Pacific. The Kuroshio enters the East China Sea with warm tropical water and exits into the North Pacific through the Tokara Strait (Figure 1), contributing to North Pacific stratification (Wang *et al.* 2016). Characterisation and monitoring of this island wake may provide important indicators of global current variations, because the Tokara Strait represents a choke point for climate cycles and change (Zhu *et al.* 2017), due to its position in the warm water transport of the Kuroshio. The Tokara Strait itself is 290 km wide, has an average depth of 634 m, and experiences an annual mean net volume transport (NVT) of $23.03 \times 10^6 \pm 3.3 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ (Zhu *et al.* 2017). The high NVT, coupled with the small size of the islands, presents ideal conditions for submesoscale island eddy trains to form.

The aim of this study is to determine the viability of identifying and studying submesoscale island wake eddies using Sentinel-3 OLCI Level 2 data, with three objectives: a) to assess the image availability provided by Sentinel-3 OLCI and its suitability for the study of submesoscale island wake formations, b) to identify submesoscale eddies and filaments, and extract chlorophyll transects of these formations for the calculation of associated metrics c) to present a novel classification framework for submesoscale eddies and filaments. It is anticipated that Sentinel-3 data will make possible the detailed study of submesoscale surface signatures, revealing a relationship between chlorophyll measurements within the features, and parameters such as current velocity, island size, and a formation's distance from the islands.

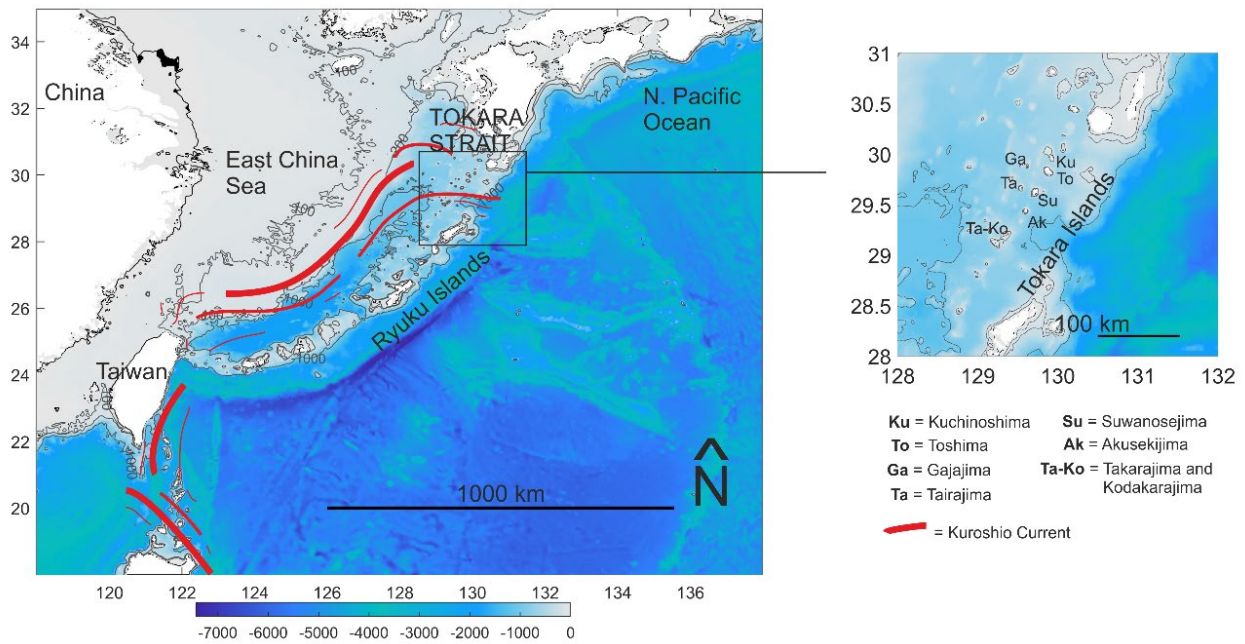


Figure 1: The East China Sea and North Pacific Ocean with subset of the Tokara Strait (ETOPO-1 bathymetry; NOAA National Geophysical Data Center. 2009: ETOPO1 1 Arc-Minute Global Relief Model. NOAA National Centers for Environmental Information. Accessed 28/5/2021; Amante & Eakins, 2009). The current field of the Kuroshio is shown in red (net transport is South to North). The study area features the Tokara Strait, within which lie the Tokara Islands (see legend) with diameters to the scale of approx. 3-10 km. The Kuroshio follows the shelf edge of the East China Sea and doglegs just before the Tokara Strait, where the flow enters the North Pacific Ocean. The velocity field of the schematic is adapted from Wang et al. (2016) and Zeng et al. (2017).

Methodology

Three types of data were acquired: multispectral ocean colour data from the OLCI onboard the EUMETSAT Sentinel-3A and -3B missions, *in situ* chlorophyll data from Biogeochemical-ARGO (Bio-ARGO) floats and Japanese Meteorological Agency (JMA) ship-based sampling, and current velocity data from Copernicus Marine Environment Monitoring Service (CMEMS) altimeter data. The phases of methodology are as follows: data procurement, data processing, algorithm validation, satellite data processing and inquisition, calculations of metrics, and statistical tests overview.

Data acquisition

The Sentinel-3 data were procured from the three available EUMETSAT data portals. All images containing pixels within the Region of Interest (ROI) were downloaded (Southwest corner: 25°N, 125°E, Northwest corner: 30.5°N, 135°E). At the time of download (November-December 2019), the individual data portals provided imaging as follows: Copernicus Online Data Access Repository (CODARep) provided data from April 2016 (Sentinel-3 mission acquisition start) until Nov 2017; EUMETSAT Datacentre provided data from December 2017 until January 2019; Copernicus Online Data Access (CODA) provided data from February 2019 until December 2019 (Table 1).

Table 1: outlines the EUMETSAT’s data portals, the imaging dates for which they provided data during the course of this study, the geographic domain (ROI) for which pixel overlap qualified a download, and the associated products that were downloaded. The products were Non Time Critical OLCI Level 2 Data (NTC OLCI L2) in netcdf format, which included all ocean colour data with associated products and coordinate grids.

Data Source	Timespan	Geographic Domain	Products
CODA	Feb 2019 – Dec 2019	25°N, 125°E - 30.5°N, 135°E	NTC OLCI L2
Datacentre	Dec 2017 – Jan 2019	25°N, 125°E - 30.5°N, 135°E	NTC OLCI L2
CODArep	Apr 2016 – Nov 2017	25°N, 125°E - 30.5°N, 135°E	NTC OLCI L3

A ‘Python Universal Sentinel Downloader’ script, provided by EUMETSAT and written by Loveday and Evers-King (2019) was used to query and download data from CODA and CODARep. The EUMETSAT Datacentre fulfils manually requested data on a case-by-case basis, so this process was not automated with a script.

In situ data were acquired from fluorometric readings of the Bio-ARGO CN-000949 and CN-000947 floats. Additional *in situ* datasets, provided by the JMA cruises from 2016, 2017 and 2018 featured chlorophyll measurements from bucket samples.

Current velocity data, for use as independent variables and potential drivers of wake characteristics, were acquired from CMEMS altimeter data and pre-processed with code developed and provided by Dr. Jill Schwarz of Plymouth University (Schwarz 2019).

Satellite (OLCI) data processing

OLCI data were processed using MATLAB in preparation for *in situ* match-ups (MATLAB 9.5.0). Grids of chlorophyll data from both the OLCI Neural Network (NN) algorithm and the OLCI_OC4Me (OC4) algorithm were re-gridded using nearest-neighbour method in order to normalise all pixel sizes to the finest possible resolution provided by the sensor (300 m x 300 m), cropped to the ROI, and filtered with quality-control flags. The ROI boundaries were selected based on requirements to maximise the available area for geographic match-ups surrounding the islands of interest, while minimising interference from coastal China waters and coastal chlorophyll plumes from Japan. The World Quality and Science Flags (WQSF) used, which are provided with all satellite data images for application as quality-control filters to eliminate bad pixels, include: ‘CLOUD’ (cloudy pixel), ‘CLOUD_AMBIGUOUS’ (partially cloudy pixel), ‘CLOUD_MARGIN’ (a margin around CLOUD and CLOUD_AMBIGUOUS of 4 pixels), ‘LAND’ (clear sky land), ‘INVALID’ (instrument data missing or invalid), ‘COSMETIC’ (missing data filled by interpolation in Level-1B), ‘SATURATED’ (saturated within any band from 400-754nm, or select other bands), ‘SUSPECT’ (potentially unreliable measures due to transmission error), ‘HISOLZEN’ (high solar zenith >70°), ‘MEGLINT’ (pixels corrected for glint), ‘SNOW_ICE’ (possible sea-ice or snow contamination), and ‘AC_FAIL’ (atmospheric correction suspect) (ESA n.d.).

Algorithm validation

Match-ups were performed in MATLAB with an in-house developed script, which specified a relatively wide match-up time window of ± 72 hours from satellite image

acquisition, in order to achieve sufficient geographic matches. (This time-window was then squeezed in post-match-up processing to determine the optimum time discrepancy for match-up accuracy). When an in-situ data point matched both the timeframe and the coordinates of a satellite pixel, the chlorophyll values for both the OLCI_NN pixel and the OLCI_OC4Me pixel were recorded. A 3x3 pixel bin average of satellite derived chlorophyll around the matched pixel was also recorded for comparison against exact single-pixel values.

The match-up process yielded no JMA data matches, and 55 Bio-ARGO data matches. All match-ups were located East of the Tokara Strait in the North Pacific Ocean (Figure 2), chlorophyll ranges differed between all data sources (Figure 3).

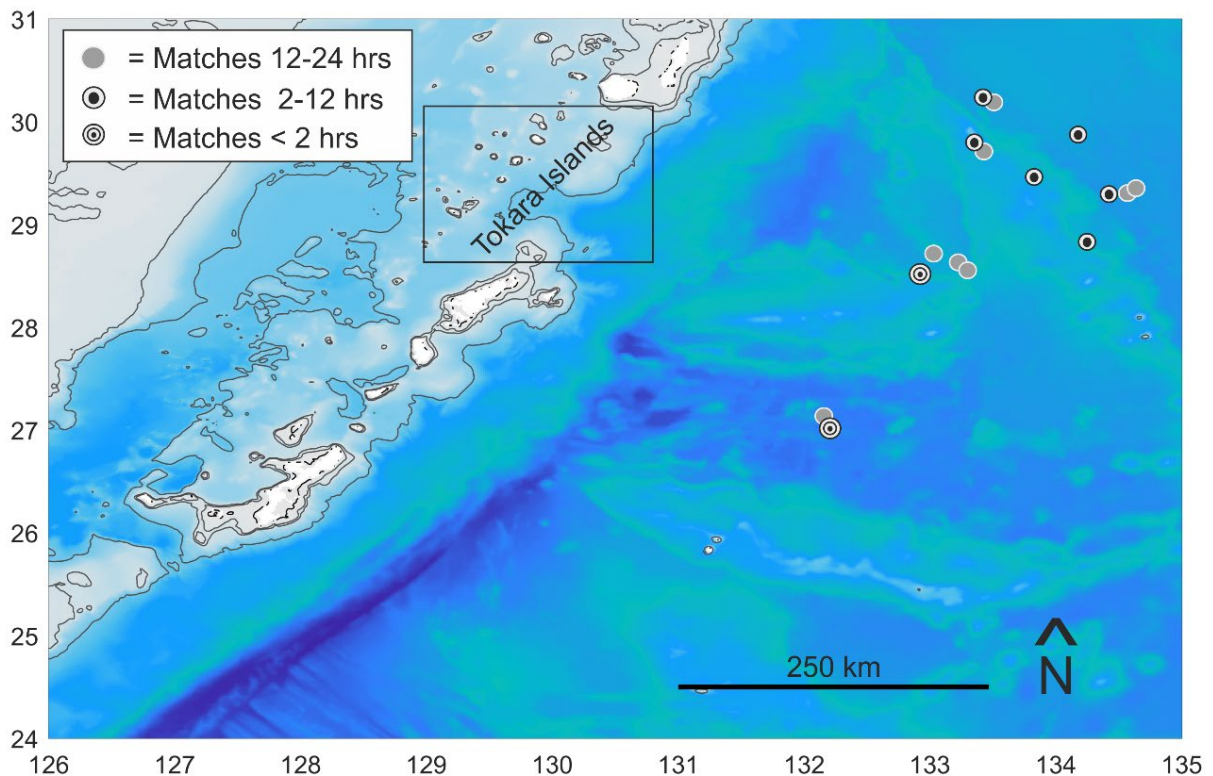


Figure 2: Map of all match-up locations within the ROI and 24-hour time window, with ETOPO-1 bathymetry as in Figure 1. Match-up markers are different for time windows of less than 2 hours, 2-12 hours, and 12-24 hours. All match-ups are located within the North Pacific Ocean, East of the Tokara Strait.

The number of matches decreased with timeframe (± 24 hours = 17 matches, ± 12 hours = 9 matches, ± 2 hours = 2 matches, with no intermediate match counts between 12 and 2 hours) (Figure 2). Type 2 linear regression was performed on data within each timeframe, in order to determine optimum match-up time threshold and most suitable algorithm for the ROI (Figure 4, (a) and (b)).

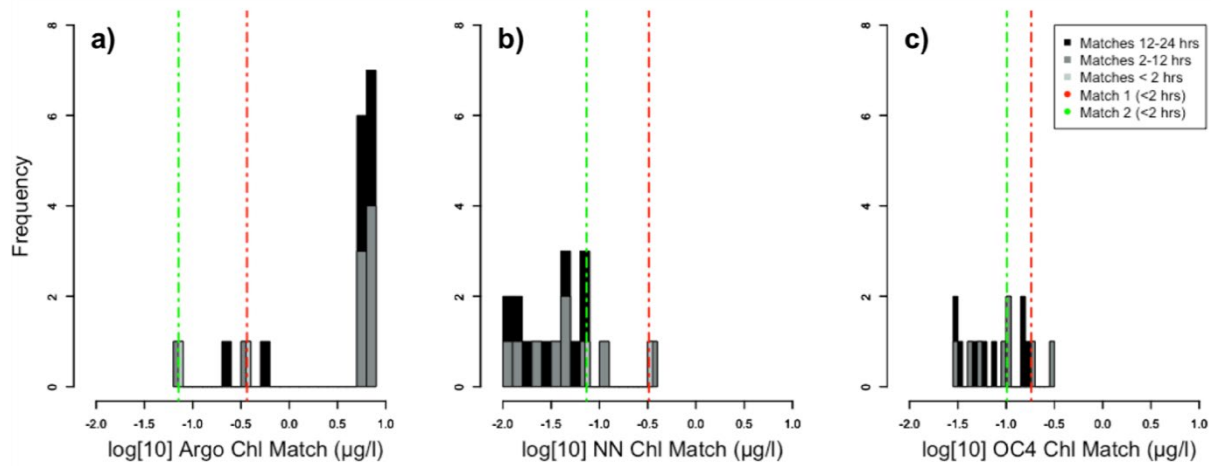


Figure 3. Histograms of the match-up values by data source. Subplot (a) is the Bio-ARGO *in situ* data, subplot (b) is NN-derived chlorophyll, and subplot (c) is the OC4-derived chlorophyll, with dotted lines marking the values of matches that fell within a 2-hour time window. The distributions of values vary between all data sources. The dotted lines denoting Match 1 and Match 2 lie around -1.1 and -0.5 for Argo and NN data, and around -1.0 and -0.7 for the OC4 data.

There was no fit within a 24 or 12-hour time window. This was attributed to the high current velocities in the region – with average velocities in excess of 0.45 m/s, water parcels can traverse up to 64 pixels in a 12-hour period. 3x3 pixel bins of NN match-ups within a 2-hour window yielded nearly 1:1 value matches with Bio-ARGO data (Table 2), and were within a half order of magnitude of each other.

Table 2: Data sources and corresponding values of both match-ups within a 2-hour time window. The values in bold, or Bio-ARGO, NN and NN 3x3 bin data present the closest matches. NN data averaged for a 3x3 pixel bin present nearly a 1:1 match to the Bio-ARGO data.

Log ₁₀ Chl (µg/l)	Match 1	Match 2
Bio-ARGO	-0.438	-1.147
NN	-0.488	-1.118
OC4Me	-0.740	-0.992
NN 3x3 Bin	-0.488	-1.134
OC4Me 3x3 Bin	-0.740	-0.992

The two matches yielded chlorophyll value ratios of 1:1.11 and 1:0.97 respectively. Although linear regression of these match-ups was not possible due to a sample size of 2, the values do suggest a linear trend (Figure 4, (c)). Considering a lack of evidence regarding one algorithm outperforming the other, the NN algorithm was chosen for this research based on reports that semi-analytical algorithms may be more robust than band-ratio algorithms (Blondeau-Patissier *et al.* 2014; Maritorena, S. *et al.* 2002).

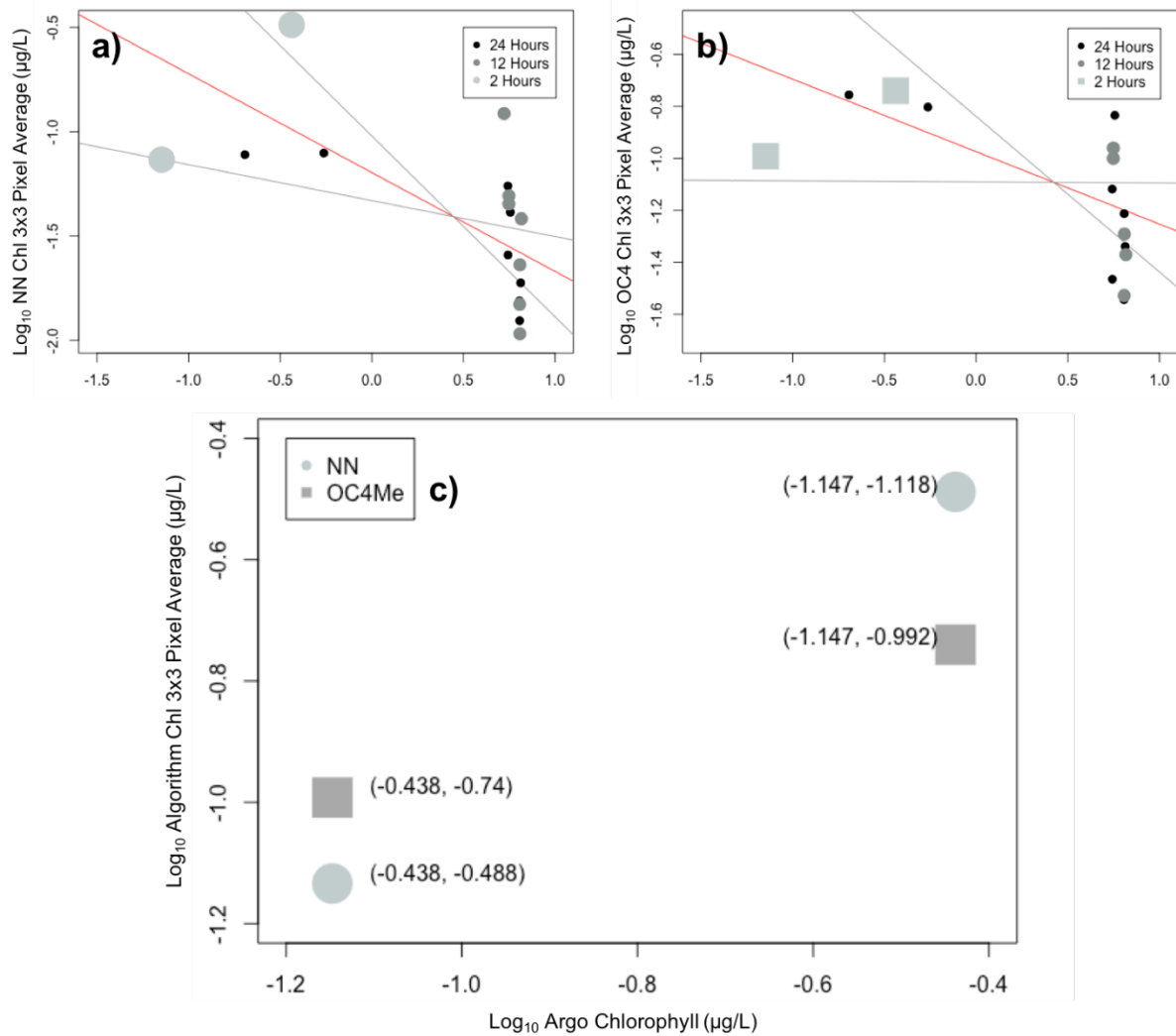


Figure 4: Linear regression analysis between 3x3 pixel bins of both CHL algorithm pixels and *in situ* CHL matches within a 24-hour period (subplots (a) and (b)). The red lines represent the line fitted by linear regression analysis. Timeframe and accuracy of match-ups are distinguished by marker colour and size respectively. Subplot (c) reduces match-ups to just the 2-hour time window, within which the NN algorithm demonstrates a nearly 1:1 match with *in situ* values, suggesting a linear relationship with intercept close to zero.

Satellite chlorophyll sampling

ESA SNAP Sentinel visualisation software was used to select images with a sufficiently cloud-clear area of interest (AOI) for sampling of the island wake. The AOI (Figure 5) is a smaller subset of the ROI (which was used only for match-ups; Figure 2), and contains only the islands and downstream wake area. In SNAP, a custom chlorophyll colormap was applied, manually adjusted for maximum contrast in the AOI, and the same quality and science flags as in match-up workflow applied (see above). The number of images per month with a clear AOI were recorded. The wake type was also recorded; a 'diffuse' wake contains no visibly identifiable eddies or filaments, whereas a wake with 'formations present' does. For each image with clear AOI, virtual sampling transects were drawn through each identifiable feature. Two transects were extracted across each eddy (one north-south and one east-west), and three transects

were extracted across each filament at manually selected spacings. Figure 5 illustrates example transects on a NN chlorophyll image visualised in SNAP, with the subset AOI expanded. See Figure 6 for the transect profile plots of the example filaments and eddies exhibited in Figure 5.

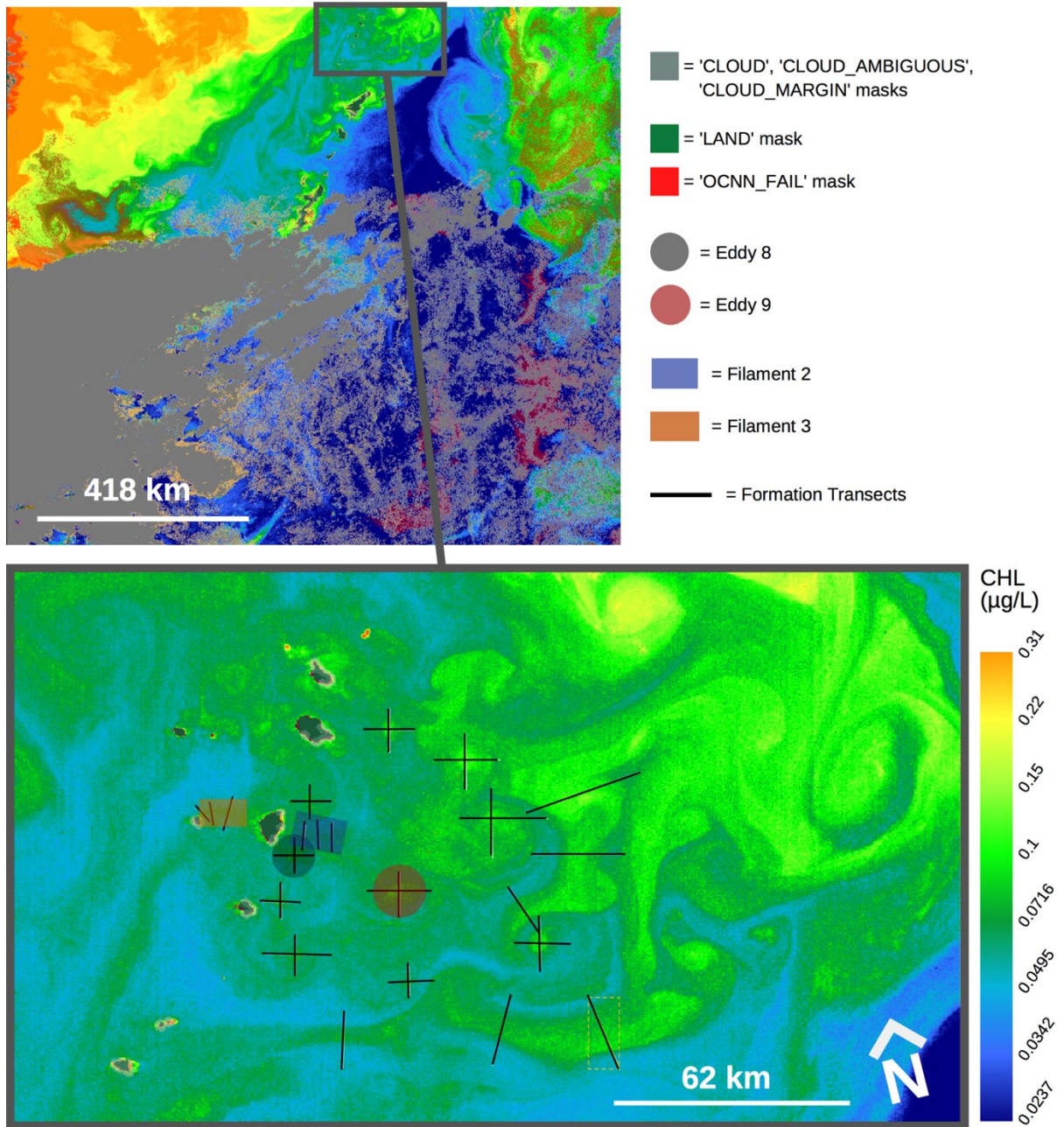


Figure 5: Initial chlorophyll image visualisation in the SNAP program, with the sampled AOI expanded. Formation transects are shown in black, with shaded areas highlighting two example filaments and two example eddies. Note that as the wake progresses to the East, interference from a chlorophyll plume (mesoscale, not addressed in this study) from the North occurs, due to larger islands in the area. No sampling was done in the area of interference.

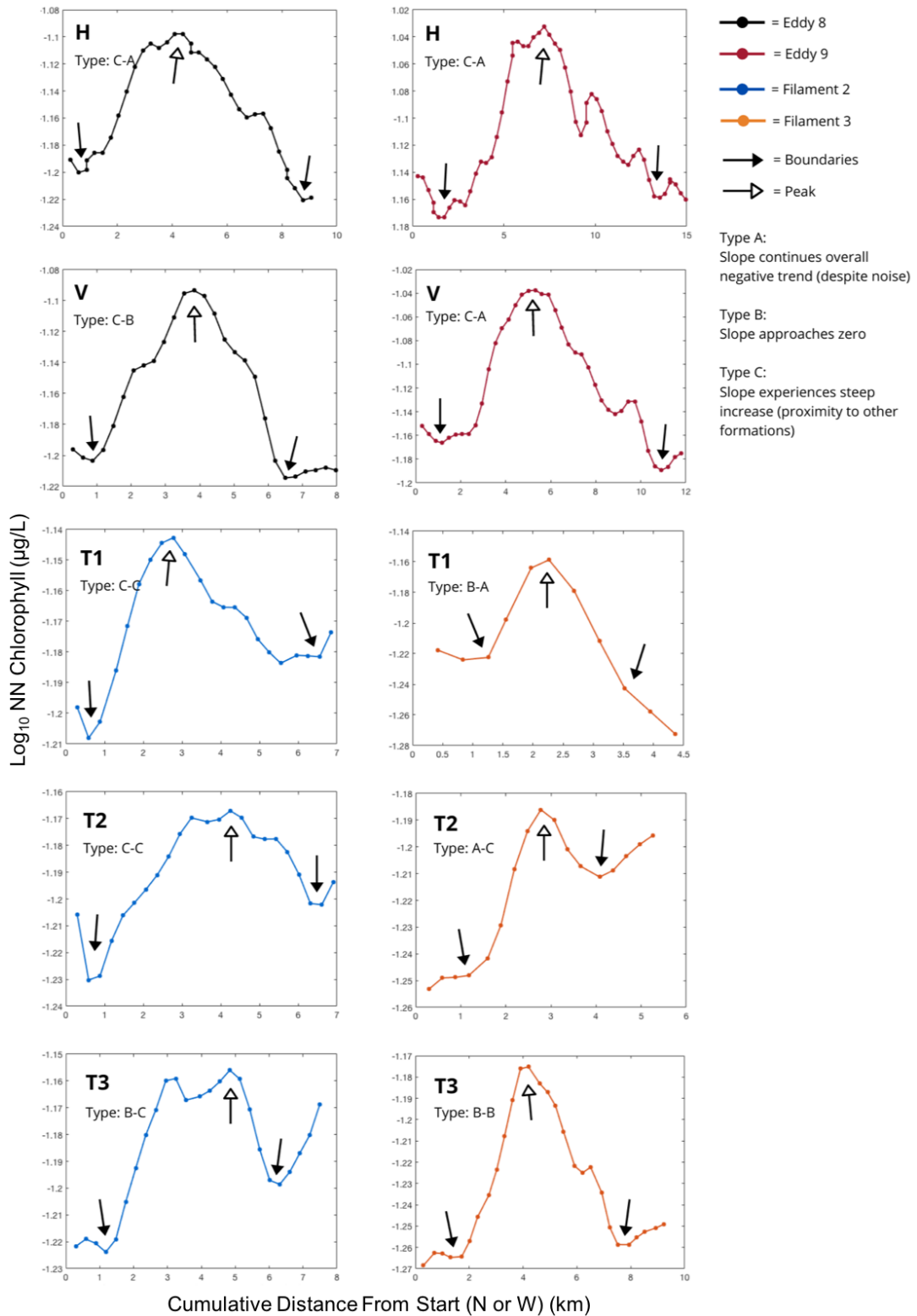






Figure 6: Subplots of all example transects taken from the data visualised in Figure 5. These illustrate transect profile types. Eddy transects are labelled (H) for horizontal east-west and (V) for vertical north-south, and filament transects are numbered T1-T3 for transects 1-3.

Calculating metrics

The peak and boundaries (Figure 6) of each transect profile were identified manually, due to the inconsistent nature of slope types within the formations. Profile slopes were classed as type A (constant decline, usually with a subtle slope change indicating the boundary), type B (slope tends toward zero at boundary), and type C (slope increases at boundary due to proximity to other formations). In some filaments, two distinct peaks were present (type P), for which the peak values and coordinates were averaged to provide a single peak for calculations (Table 3). Figure 6 shows most profile slope types (A through C), though lacks a type P filament, which were scarce.

Table 3: Profile types alongside visual and written descriptions of the slope characteristics that define each type and guide manual selection of the peaks and boundaries.

Profile Type	Example Schematic	Description
A		Slope from peak to end stays negative, usually with subtle slope change indicating boundary.
B		Slope tends toward zero at boundary (reaching ambient island wake concentrations).
C		Slope becomes positive at boundary, due to proximity to other chlorophyll maxima-inducing formations.
P		Double-peaked filament: two distinct peaks present, usually separated by clear drop in concentration at centre.

After recording peak and boundary values, the formation metrics (dependent variables) were calculated. These comprised chlorophyll gradient as a measure of chlorophyll concentration strength, and full width at half max (FWHM) of the profile peak as a measure of formation size. Gradient was calculated with the following broad formula:

$$G = \frac{\delta y}{\delta x}$$

(Equation 1)

where G is the average gradient of both sides of the profile, δy is the average chlorophyll amplitude, and δx is the average distance between boundary 1 and the centre of the curve, and boundary 2 and the centre of the curve, as determined in this equation:

$$G = \frac{\left(\frac{(c_2 - c_1)}{(d_2 - d_1)} + \frac{(c_2 - c_3)}{(d_3 - d_2)} \right)}{2}$$

(Equation 2)

where G is the overall gradient of the formation, c_1 is the boundary 1 chlorophyll, d_1 is the boundary 1 cumulative distance, c_2 is the peak chlorophyll, d_2 is the peak cumulative distance, c_3 is the boundary 2 chlorophyll, and d_3 is the boundary 2 cumulative distance.

The chlorophyll value at FWHM was calculated as follows:

$$C = c_2 - \left(\frac{c_2 - \left(\frac{c_1 + c_3}{2} \right)}{2} \right)$$

(Equation 3)

where C is the chlorophyll value at FWHM, c_1 is the first base chlorophyll value, c_2 is the chlorophyll value at the peak, and c_3 is the second base chlorophyll value. The expression in the numerator calculates the amplitude of the profile; the two baseline values are averaged to achieve a unified baseline. Once the chlorophyll value at FWHM was identified, the cumulative distance was interpolated to the closest possible location of the FWHM chlorophyll values in the profile. The distance between these locations (cumulative distance at HM point 1 subtracted from cumulative distance at HM point 2) is the FWHM (see Figure 7).

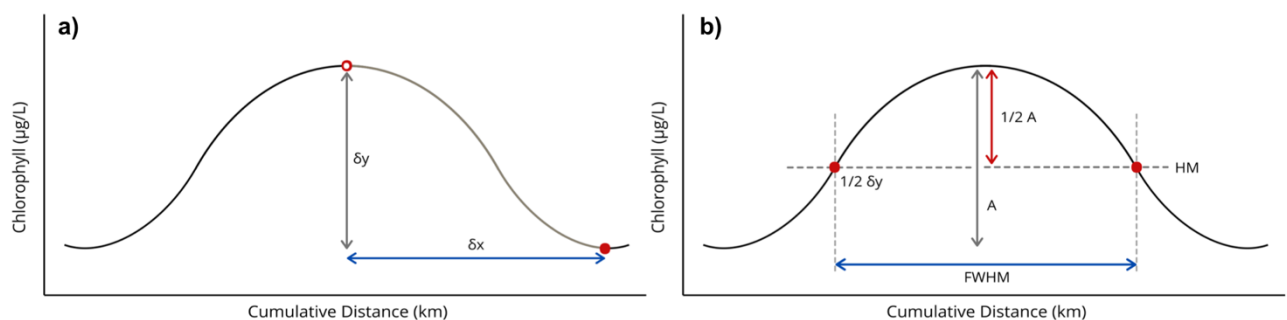


Figure 7: Visual representation for the calculations of formation metrics. Subplot (a) shows the vertical (chlorophyll concentration) and horizontal (kilometres) distances used to calculate gradient, and subplot (b) shows the further dimensions used to calculate FWHM.

All gradient values were categorised to provide a biologically relevant classification framework for submesoscale formations. Gradient levels proposed herein are defined by the 25th and 75th percentiles of all gradient values. Weak gradients are within the 25th percentile (> -2.481), moderate gradients are between the 25th and 75th percentile (< -2.481 and > -2.153), and strong gradients are above the 75th percentile (< -2.153) (Table 4; Figure 8).

Table 4: Summary of chlorophyll gradient classes, as defined by the 25th and 75th percentiles and associated chlorophyll value ranges.

Gradient Class	Percentile Range	Value Range
Weak	< 25 th	-2.481 – -2.962
Moderate	> 25 th & < 75 th	-2.291 – -2.481
Strong	> 75 th	-0.599 – -2.291

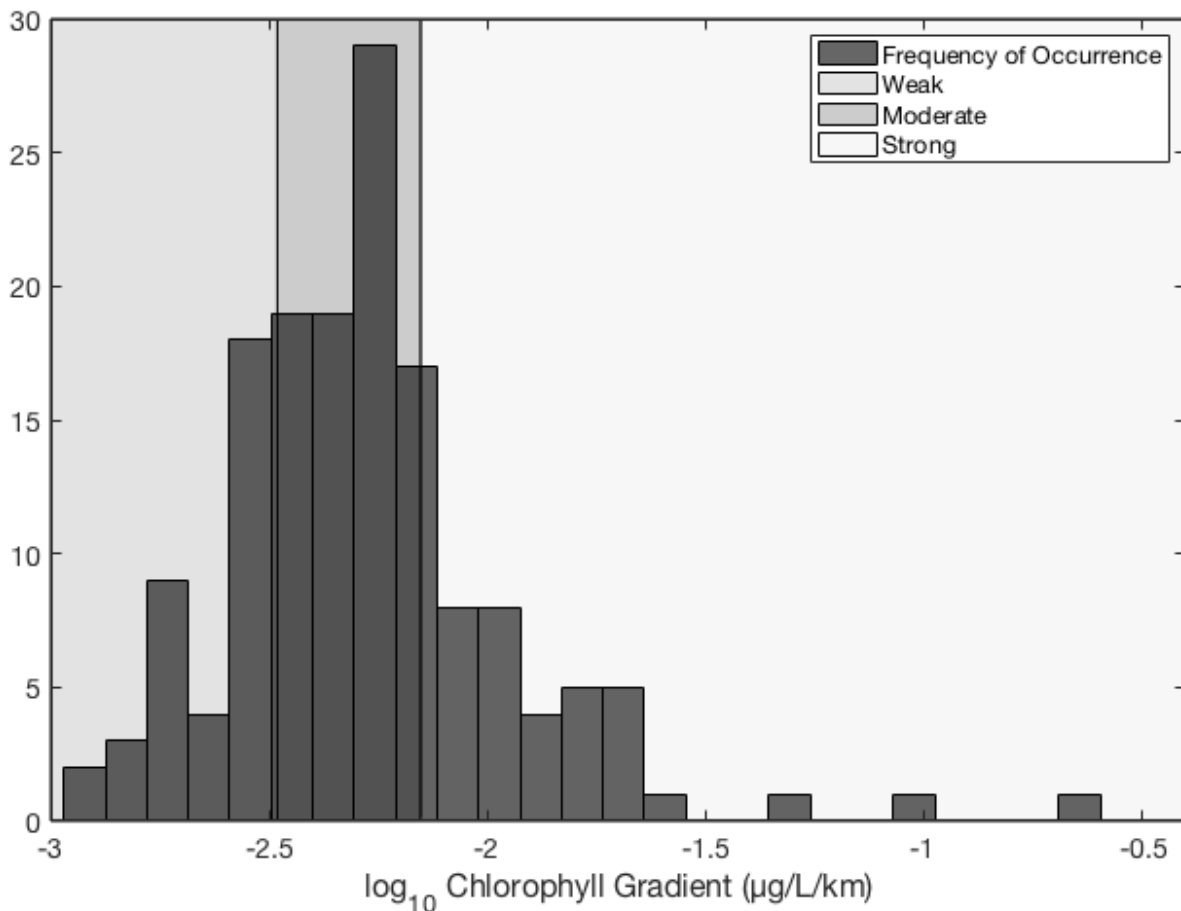


Figure 8: Histogram of all chlorophyll gradient values, with shaded regions highlighting the classification levels as defined by the 25th and 75th percentiles.

All calculated metrics were then analysed against the environmental variables: distance from island, island diameter, month, current speed and current angle. These variables were cross analysed for interactions before statistical tests were undertaken in R. Multiple linear analysis (MLA) models were used to analyse all continuous variables against gradient and FWHM. This was done for eddy data and filament data independently, as these present distinct formation types which may have varying responses to environmental drivers.

Results

Data availability for AOI

The viability of small-AOI study varies seasonally. The average number of total OLCI images available varied monthly (Table 5), though all fell within a range of N = 20 to 30. EUMETSAT image quality checks eliminated some images from processing, while other variability in image availability is attributable to differences in data search capabilities between the different data sources.

Cloud cover at the time of image acquisition was the greatest influence on monthly variability in the usable proportion of images. The percentage of images with a clear AOI exhibits a seasonal pattern (Figure 9, (a)), with maximum clarity during midsummer, between months 5 and 7. Months 1, 2, 3, and 12 provide only 0.91 – 3.33% chance of finding an image with clear AOI (Table 5).

No eddies or filaments were identifiable during December, January, and February for all years (Figure 9, (b)). This may be due solely to AOI clarity inhibited by cloud cover in the region (Table 5) however the diffusivity of the island wake exhibits a seasonal pattern that is inverse to that of formation identification, so this may also contribute. Figure 9 (a) shows that the percentage of diffuse island wakes, out of all visible wakes for each month, reaches 100% for December, January and February. It is important to note, however, that the sample size during these months establishes a dataset-wide minimum (Table 5).

Table 5: Monthly data usability metrics and formation counts. Image clarity, wake type, and formation counts exhibit seasonal variability; winter months provide reduced clear images, and an increased percentage of diffuse wakes, thus reducing eddy and filament counts. Image clarity for the area of interest is expressed as a percentage of the average images available each month. The ‘Diffuse Wakes’ column expresses the number of wakes within which no formations were identifiable as a percentage of the total images with clear AOI.

Month	Images Available	Images w/ Clear AOI (%)	Diffuse Wakes (%)	Avg Eddies	Avg Filaments
1	21	1.59%	100.00%	0	0
2	20	3.33%	100.00%	0	0
3	23.33	2.86%	33.33%	0	0.33
4	28	7.14%	20.00%	2	2.25
5	26	7.69%	42.86%	2	2.75
6	25	8.00%	25.00%	0.75	0.5
7	26.25	7.62%	12.50%	2	2.75
8	26.5	6.60%	0.00%	4.75	3
9	25.25	5.94%	28.57%	5.5	0
10	25.75	5.83%	0.00%	5.25	1.25
11	24.5	5.10%	40.00%	3	1
12	27.5	0.91%	100.00%	0	0

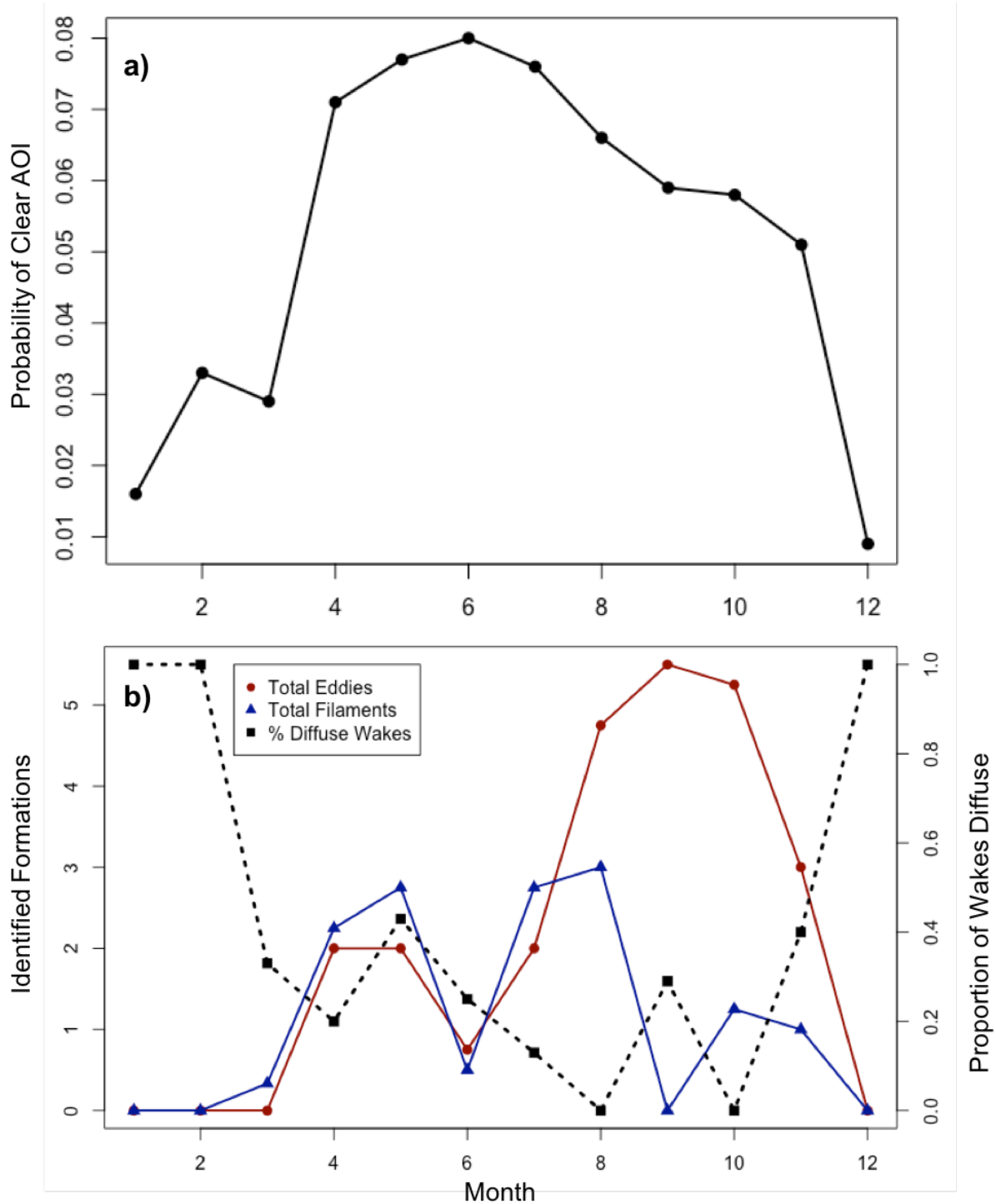


Figure 9: Clear-AOI imagery variability (subplot (a)) and wake type and formation count variability (subplot (b)) per month. Both the image clarity and the number of identified formations are low in winter months, when diffuse wakes prevail. For quantification of these patterns, see Table 4.

Monthly-seasonal formation presence

The total number of identified eddies exhibits an overall positive trend from mid-spring (month 4) until early autumn (month 10), peaking in month 9 (Figure 9 (b)). The total number of identified filaments follows a gently inverse trend, with higher counts present between months 4 and 8, and decreasing thereafter (Figure 9 (b)). The

percent of diffuse wakes follows the opposite tendency to that of identifiable formation presence, with the highest proportion of diffuse eddies (exhibiting no visually distinct / identifiable formations) occurring when formation counts were absent or minimal. The smallest proportion of diffuse wakes occurred simultaneously to the highest number of eddies, between months 8 and 10 (Figure 9 (b)).

Formation gradients

While total eddy and filament counts showed peaks in different months, so did counts when broken down into gradient level categories of weak, moderate and strong. Weak-gradient eddies peak in month 8 with an average count of 2.75, moderate-gradient eddies peak in month 10 with an average count of 5, and strong-gradient eddies peak in month 4 with an average count of 1.75 (Table 6; Figure 10). For filament counts, peak months for strong, moderate and weak groups differ from those of eddies. Weak filaments peak in months 10 and 11 with average counts of 0.75, moderate filaments peak in month 8 with count average of 1.25, and strong filaments peak in month 5 with count average of 1.5 (Table 6; Figure 10).

Table 6: The monthly occurrence of weak, moderate and strong gradients within eddies and filaments. The presented values are monthly means across 3-4 years of data, depending on the month.

Month	Weak Eddies	Moderate Eddies	Strong Eddies	Weak Filaments	Moderate Filaments	Strong Filaments
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0.33	0
4	0	0.25	1.75	0	0.25	2
5	0.5	0.5	1	0.5	0.75	1.5
6	0.25	0.5	0	0	0.5	0
7	0.5	1.5	0	1	1.5	0.25
8	2.75	1.5	0.5	1	1.25	0.5
9	1.75	2.25	1.5	0	0	0
10	0	5	0.25	0.75	0.5	0
11	0	2.5	0.5	0.75	0.25	0
12	0	0	0	0	0	0

The gradient data variation can also be summarised into seasonal patterns of occurrence. During northern hemisphere spring, strong formations were the most prevalent for both filaments and eddies (Figure 10). In mid-late summer, weak eddies become the most prevalent of their formation type, while moderate filaments surpass both weak and strong filament counts. The occurrence of moderate eddies peak in autumn months, far surpassing both weak and strong eddies. In these same months, weak filaments are most dominant (Figure 9). The seasonal variability can be summarised as: winter has no detectable structure; spring has strong eddies and strong filaments; summer has weak eddies and moderate filaments; autumn has moderate eddies and weak filaments.

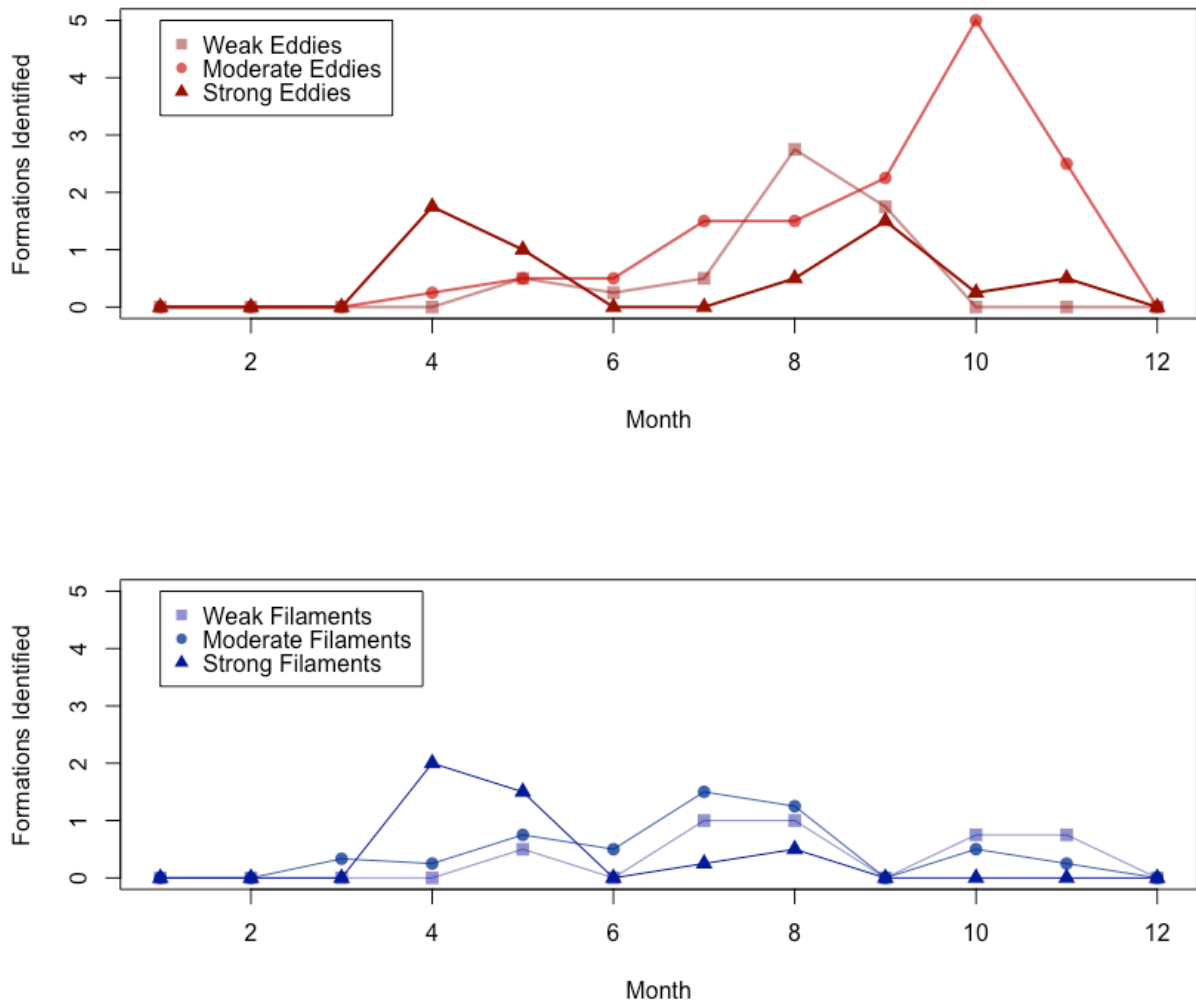


Figure 10: The average number of eddies and filaments identified per month. Eddies are plotted in shades of red, and filaments are plotted in shades of blue. Strong gradient counts are marked by triangles, moderate by circles, and weak by squares, with colour saturation scaling with gradient for ease of visualisation. Strong eddies peak in spring, weak in summer, and moderate in autumn. Strong filaments also peak in spring, though moderate filaments peak in summer, and weak filaments in autumn.

Contributing variables

Upon screening independent variables for interactions using analysis of variance (ANOVA), current speed was found to vary significantly with month (ANOVA; p -value $< 1 \times 10^{-16}$, F -value = 58.81). Current speed is retained as the variable of choice rather than month because it allows incorporation of all continuous variables into a multiple linear regression (MLR) analysis. Figure 11 shows peak speeds in month 4 (corresponding with dominance of high filaments and eddies; Figure 10) decreasing steadily until month 11. Most notably, months 4 and 5 varied greatly from all other months (Tukey HSD; p -values < 0.05), though months 6 through 11 did not show significant variance amongst each other (Tukey HSD; p -values > 0.05). Because current speed is recorded parallel to formation identifications, data is lacking for winter months.

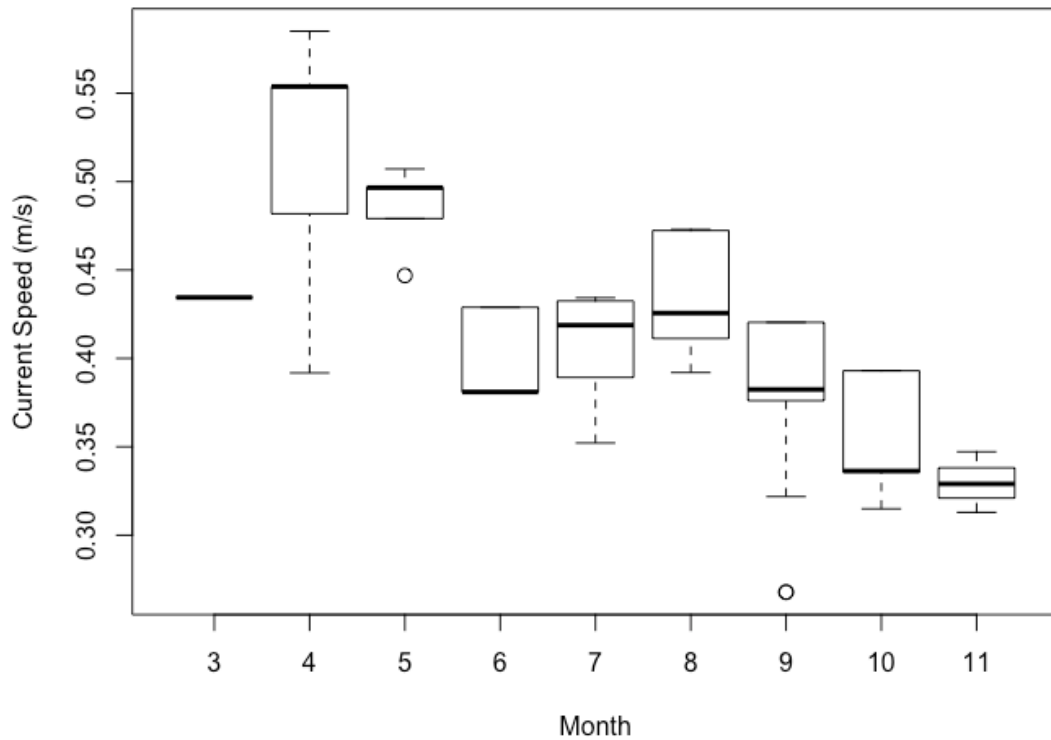


Figure 11: Current speed on successful AOI imagery days, per month. Current speed decreases from April to November. The peak current speeds fall in spring, aligning with strong eddy and filament gradients. No data is shown for December – February when no clear imaging was available.

Correlative variables

Using MLR models, all continuous variables (current speed, current direction, distance from island, and island diameter) were tested for their correlation with a) gradient, and b) FWHM. The MLR of eddy gradient (MLR; p-value = < 0.001, Residual Standard Error = 0.286 on 96 DF, and R-squared = 0.219) (Table 7) yielded two significant independent variables: current speed (MLR; slope = 1.490, p-value = 0.001) and current direction (MLR; slope = 1.660×10^{-02} , p-value = 0.016) (Figure 12, subplots (a) and (b)).

Table 7: The multiple linear regression statistical analyses results. MLR were conducted between all continuous environmental variables (distance from island, island diameter, current direction, and current speed), and both calculated eddy and filament metrics (gradient and FWHM). Note the P-value, Multiple R² value and Residual Standard Error in the upper right corner of each subtable.

Eddies MLR Descriptive Statistics / GRADIENT			Eddies MLR Descriptive Statistics / FWHM		
P-value = **7.972 x 10 ⁻⁰⁵ Multiple R ² = 0.219 Residual Std Error = 0.286 on 96 DF			P-value = **1.285 x 10 ⁻⁰⁵ Multiple R ² = 0.250 Residual Std Error = 1.711 on 96 DF		
Independent Variables:	Slope	P-value	Independent Variables:	Slope	P-value
Distance from Island	1.916 x 10 ⁻⁰⁵	0.986	Distance from Island	0.015	*0.024
Island Diameter	-1.495 x 10 ⁻⁰²	0.330	Island Diameter	0.361	**< 0.001
Current Direction	1.660 x 10 ⁻⁰²	*0.016	Current Direction	-0.047	0.243
Current Speed	1.490	**0.001	Current Speed	10.205	**< 0.001

Filaments MLR Descriptive Statistics / GRADIENT			Filaments MLR Descriptive Statistics / FWHM		
P-value = **< 0.001 Multiple R ² = 0.317 Residual Std Error = 0.311 on 49 DF			P-value = 0.183 Multiple R ² = 0.117 Residual Std Error = 2.111 on 49 DF		
Independent Variables:	Slope	P-value	Independent Variables:	Slope	P-value
Distance from Island	-0.007	*0.016	Distance from Island	0.030	0.110
Island Diameter	0.002	0.944	Island Diameter	0.231	0.118
Current Direction	0.003	0.713	Current Direction	0.056	0.371
Current Speed	2.686	**< 0.001	Current Speed	0.863	0.863

Because these two trends appear similar, speed and direction were analysed for interference (Pearson’s Correlation; 0.293), revealing 30% correlation between the two elements of velocity. The MLR of eddy FWHM (MLR; p-value = < 0.001, Residual Standard Error = 1.711 on 96 DF, and R-squared = 0.250) (Table 7) yielded three significant independent variables: formation distance from island (MLR; slope = 0.015, p-value = 0.024), island diameter (MLR; slope = 0.361, p-value = < 0.001), and current velocity (MLR; slope = 10.205, p-value = < 0.001) (Figure 12, subplots (c), (d) and (e)).

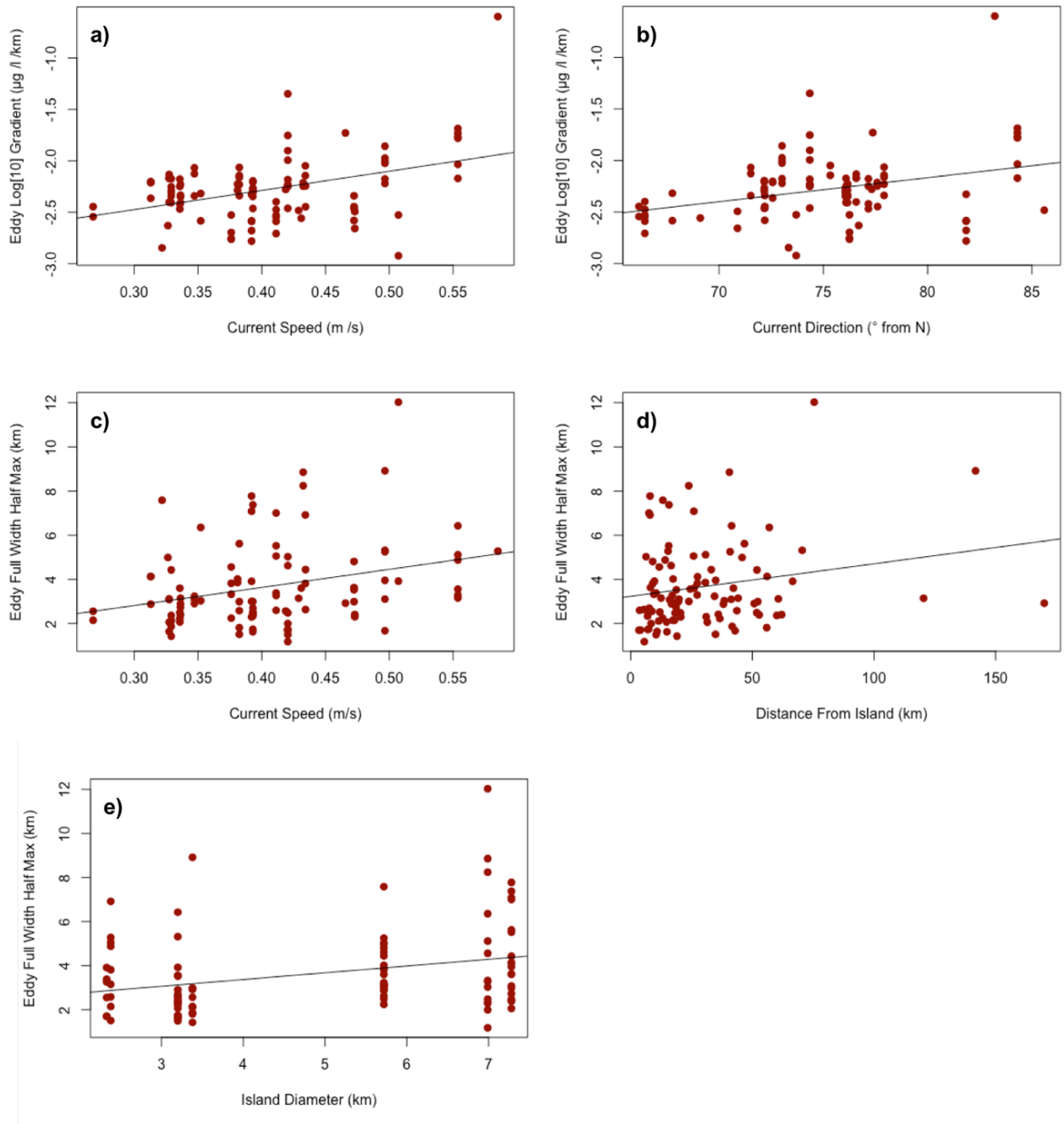


Figure 12. The relationships between eddy gradient and all significantly correlated variables, and eddy FWHM and all significantly correlated variables. Current speed and direction are both significantly positively correlated with eddy gradient (subplots (a) and (b)), and current speed, distance from island and island diameter are all positively correlated with FWHM (subplots (c) – (e)).

MLR of filament gradient (MLR; p -value = < 0.001, Residual Standard Error = 0.311 on 49 DF, and R-squared = 0.317) (Table 7) yielded two significant correlative variables: current velocity (MLR; slope = 2.686, p -value = < 0.001) (Figure 13, subplot (a)) and distance from island (MLR; slope = -0.007, p -value = 0.016) (Figure 13, subplot (b)). Like eddy gradient, filament gradient shows positive correlation with current velocity. However, MLR of filament FWHM was not statistically significant.

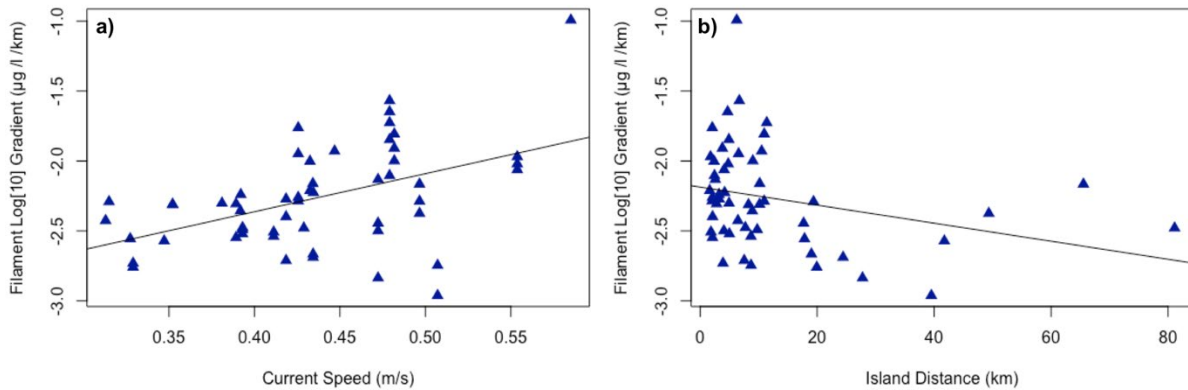


Figure 13: The relationships between filament gradient and all significantly correlated variables. Current speed is significantly positively correlated with filament gradient (subplot (a)), and distance from island is significantly negatively correlated with gradient, implying that filaments become less strong as distance progresses.

Discussion

This research achieves three outcomes: a) evaluates the viability of using the most advanced ocean colour imagery currently available to study submesoscale island wake features, b) corroborates previous studies' conclusions regarding the environmental variables driving island wake characteristics and provides new insight, and c) expands current knowledge of submesoscale island wakes from a physical to a biological interpretation. The discussion below is partitioned into a thorough examination of each of these individual outcomes.

Imaging success and limitations

By using the highest spatiotemporal resolution ocean colour imagery available, this study begins to overcome previous limitations regarding coarse temporal frequency of re-imaging. However, this is still subject to extensive limitations that are seemingly not solved by higher temporal imaging resolution. Although ocean colour imagery from the LANDSAT satellites provides spatial resolution on par with Sentinel-3 (30 m over ocean), it has coarse temporal resolution with re-imaging time of around 12 days (Isoguchi *et al.* 2009). Limited temporal frequency of imaging makes it difficult to achieve statistically significant results regarding intra- and inter-seasonal variations and, given the transient nature of submesoscale formations, renders impossible the monitoring of real-time progression of island wake formations. Given current velocities measured in the Tokara strait during this research period of between 0.32 and 0.55 m/s, a parcel of water can move by up to 66 pixels per hour. During a 12-day period, the centre of an eddy could therefore move by up to 190,008 pixels or 6,333 m. Considering this spatial progression along with submesoscale temporal evolution timescales of 1 day (Hosegood, 2017; Mahadevan, 2006), any transient formation

would be impossible to monitor in real time. Submesoscale island wakes therefore present fast-developing and highly variable study areas, adding to the existing challenge of achieving frequent cloud-free imagery.

In contrast to LANDSAT, Sentinel-3A and -3B OLCI imagery offers much finer temporal sampling, providing repeat imaging twice per day. Despite initial optimism about the viability of leveraging this improvement to study the small AOI by reducing the recurrence interval of cloud-free imagery, the potential of these data for submesoscale research remains limited. Even given frequent repeat imaging, ocean colour data remains inherently constrained to daylight hours and cloud-free zones. This study found that cloud cover reduced the number of usable images for the AOI by over 90% each month, preventing the ability to monitor the real-time progression of specific eddies and filaments within the island wake. Despite presenting a great advancement in temporal sampling resolution over LANDSAT, cloud cover still limits the OLCI to providing only isolated snapshots in time. Inventory of usable images showed a seasonal pattern, with the probability of achieving clear AOI being greatly improved in northern hemisphere summertime. Cloud cover over a small AOI and a sampling bias due to seasonality of cloud cover continue to challenge the viability of using these data for submesoscale research. Although monitoring the progression of submesoscale formations remains infeasible, classification of the eddies and filaments using a novel gradient-based approach was still possible.

Physical drivers of island wake characteristics

In contrast to the present use of the spatial distribution of chlorophyll as the primary indicator of turbulent structure, previous remote sensing studies of submesoscale island wake formations have averted the limiting influence of viable image availability by using synthetic aperture radar (SAR), or infrared (IR) bands for sea surface temperature (SST), sometimes coupled with physical models (Isoguchi *et al.* 2009; Zheng *et al.* 2008; Zheng and Zheng 2014). Neither of these imaging techniques are limited by light availability and are less impeded by cloud cover, therefore providing generally finer temporal resolution. It is important to keep in mind that these studies are limited to a detailed understanding of the physical features, as they do not examine an associated biological signature (chlorophyll). These studies nonetheless provide invaluable information on the physical behaviour and characteristics of island wake formations. They provide the foundation for this research and present evidence regarding the driving factors and potential importance of these turbulent island wake formations.

The SAR and IR based studies offer corroborating evidence regarding driving variables of island wake formations. Specifically, these studies corroborate the findings herein that wake characteristics vary with current velocity. In 2014, Zheng and Zheng (2014) used the European Remote Sensing (ERS-1) satellite SAR imagery to look at spatial characteristics of island wake vortex trains in the lee of Orchid Island and Green Island, Taiwan. They found that the eddies were current-induced rather than wind-induced. Later, Hsu *et al.* (2019) saw that current velocity determines the characteristics of the wake. Without sufficient speed, and without the main axis being intercepted by the island, no wake would occur (Hsu *et al.* 2019). The present research further revealed that current velocity influences not only the entire wake type, but the structure of individual eddies and filament.

The present results show a correlation between current velocity and eddy gradient, and also a correlation between current velocity and month. This evidence strongly implies the existence of systematic seasonal variation in eddy and filament gradient types. This is probably due to seasonal variations in the strength and main axis shift of the Kuroshio. Liu *et al.* (2019) observed that in the Tokara Strait, the Kuroshio experiences a northward shift coupled with increased current velocity in January, February and July, and a southward shift coupled with increased current velocity in December. They found the main axis location to be most variable throughout June and September. Despite all this variability, the dominating trend is that currents are strongest in July, and weakest in October (Liu *et al.* 2019). Although wake characterisation capabilities for the purposes of this research were reduced during winter months due to cloud cover, thus limiting conclusions regarding the effect of Dec-Feb intensification, the spring, summer and autumn months present a trend which may be related to the July through October current weakening.

The missing link

The eddy-denoting chlorophyll signatures used in this study may provide the missing data link between the postulated importance of these turbulent formations, and the greater ecological implications of such. Although previous studies have inferred the importance of the island wakes on surrounding ecosystems (Isoguchi *et al.* 2009; Liu *et al.* 2019; Hsu *et al.* 2019; Zheng and Zheng 2014), none have used a biological indicator (e.g., chlorophyll) as the primary measurement of turbulent mixing formations with potential ecological importance. Some, however, looked briefly at chlorophyll imagery in order to point out the potential biological implications of turbulent island wake formations, thus giving this study precedent to use chlorophyll as a primary indicator of turbulent formations. Isoguchi *et al.* (2009) first used the Advanced Land Observing Satellite (ALOS) Phase-Array L-band Synthetic Aperture Radar (PALSAR) to identify a 30 km – long island wake in the lee of the Izu Islands, Japan, and then used LANDSAT SST and CHL images to look at features on a finer scale. Although the SST and CHL data were not used to monitor or characterise the features, they did confirm that mixing and productivity associated with island-mass effect was occurring. Hsu *et al.* (2019) confirmed evidence of strong upwelling based on significantly colder eddy cores within submesoscale island wake vortices.

While using biological indicators, this study's gradient-based characterisation method also provides a finer scale examination of individual wake features than previous island wake characterisations. Hsu *et al.* (2019) presented a comparable method of characterising island wakes using SST imagery, characterising wakes observed from Green Island into 4 types based on how far downstream the tail of the wake stretched, and whether vortices were visibly present. The classification developed in the present research identifies eddies and filaments based on their chlorophyll gradient value, providing a method by which to group individual formations by their biological signature.

A study by Zhang *et al.* (2019), which was published while this research was underway, presents a novel inquiry of submesoscale eddies against which to compare the methods of this paper. Although the eddies queried by Zhang *et al.* (2019) were associated with a larger mesoscale circulation system in the Florida Straits and not island wakes, this study presents a successful elucidation of submesoscale turbulent structure using chlorophyll image products. Zhang *et al.* (2019) used novel (custom developed) MODIS ocean colour products and modified algorithms, and machine

learning to interpolate and recover low-quality or corrected chlorophyll data to overcome limitations of available NASA products that are filtered by sun-glint and cloud cover. The Sentinel-3 data products used for the research presented here are also subject to such filters, severely depleting the available data. This is the primary limitation to achieving continuous monitoring of submesoscale eddy wakes, and a solution to this issue would be most advantageous. However, as of yet there is no confirmation of the accuracy of the interpolative methods employed by Zhang *et al* (2019), and this would require much fine-tuning for an area as concentrated and transient as an island wake.

Furthermore, Zhang *et al.* (2019) used a canny edge detection algorithm to detect eddies, for which they determined the optimum boundary thresholds through trial-and-error of 625 possible values. These optimum values are specific to the eddies of interest in the area of interest, although the vague application of this method may be suitable for island wake eddies, potentially improving upon the method used in this study by reducing manual data processing time, however this must pend confirmation that the algorithm provides equal or superior eddy detection accuracy to that of human image interpretation. Potential advantages of using the manual methods in this paper rather than the canny edge detection include: ability to identify filaments and not just circular formations, and the ability to more generally characterize each island wake based on the level of diffusivity. Furthermore, the manual wake formation sampling method presented here is easily accessible to any biological oceanography-oriented scientist with an interest in examining island wake characteristics, without the need for advanced custom algorithms.

Biological implications and new directions

This research presents a foundation on which to begin coupling surface chlorophyll signatures of submesoscale island wakes with exploitation by higher trophic level organisms. Oceanographic features such as upwelling, wind-driven mixing, internal waves, tidal fronts, and other aggregating processes dictate dynamics not just of phytoplankton, but all trophic levels – from primary, to secondary, to tertiary producers – inferring that the great importance of these features can be extended to those discussed herein. The effect of gradients on the microdiversity, relative abundance and niche differentiation of microorganisms such as cyanobacteria is well documented (Dang *et al.* 2019). For primary and secondary producers, turbulence and upwelling dictate nutrient availability, light limitation, and prey capture rates (MacKenzie 2000; Rothschild 1988). Environmental and biological gradients are known to host elevated concentrations of ichthyoplankton (Holliday 2011; Atwood 2010), and may be crucial foraging hotspots for pelagic fish, seabirds, and harbour porpoises (Scales 2014; Godo 2012; Jones 2014). The biological implications listed here may also be applicable to submesoscale island wake formations, making them of potentially great importance to oceanic ecological processes.

The weak/moderate/strong chlorophyll gradient classes presented in this paper were chosen as a metric rather than absolute chlorophyll concentration because of the importance of gradients to ecological processes in the oceans. Given the necessity for marine predators to detect prey aggregations, one might predict that deviations from mean ambient chlorophyll concentration in an AOI will be more important to detection and subsequent exploitation of chlorophyll-related production hotspots than will absolute concentration within the eddy or filament. This classification framework, when applied to specific conditions in any area of interest, might be used to identify any

differential exploitation of eddy or filament classes by marine predators throughout the trophic levels. Pending further study, classifying filaments and eddies based on gradient may be a proxy for biological significance of the turbulence structure. For further research, a question remains: considering time lags between chlorophyll blooms and subsequent foraging and predation movements, are submesoscale island wake formations persistent enough to produce reliable foraging habitat for higher trophic levels?

Continuation of this research may be used to inform monitoring and management efforts of small island fisheries in the face of global climate change. As exhibited by this and previous research, island wake systems are immediately sensitive to environmental variations. As mentioned by Zhu *et al.* (2017), the Tokara Strait in particular may present a location where climate change effects on currents and associated processes become evident. Island wake turbulent systems may therefore present critical locations at which to monitor the direct physical and biological implications of shifting global circulation patterns. Considering that island communities are some of the most vulnerable to climate change (Cicin-Sain 2015), understanding the sensitive dynamic processes and associated biological implications within artisanal island fisheries is imperative to sustainable management. The methodology developed in this research can help contribute to such an understanding.

Acknowledgements

The author acknowledges the great guidance and mentorship of Dr Jill Schwarz. Only with her patience, faith, teaching and expertise was this research made possible. Many thanks also to Dr Clare Embling, for her constant support and engagement in the success of her tutees. These two role models were instrumental in the author's ability to culminate her university career in the production of this research.

References

Amante, C and B.W. Eakins, 2009. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. doi:10.7289/V5C8276M [28/5/2021].

Antoine, D. OLCI Level 2 algorithm theoretical basis document: Ocean colour products in Case 1 waters. N/A: 2010.

Atwood, E. et al. (2010), 'Influence of mesoscale eddies on ichthyoplankton assemblages in the Gulf of Alaska', *Fisheries Oceanography*, 19(6), pp. 493-507.

Blondeau-Patissier, D. et al. (2014), 'A review of ocean colour remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans', *Progress in Oceanography*, 123, pp. 123-144.

Chen, X. (2004) Analysis of the circulation on the East-Chinese Shelf and adjacent Pacific Ocean. PhD Thesis. University of Hamburg. Available at: <https://www.researchgate.net/publication/33957506> (Accessed: 3 Feb 2020).

Cherubin, L. and Garavelli, L. (2016) 'Eastern Caribbean circulation and island mass effect on St. Croix, US Virgin Islands: A mechanism for relatively consistent recruitment patterns', *PLoS One*, 11(3), pp. 1-28.

Cicin-Sain, B. (2015) *Goal 14 – conserve and sustainably use oceans, seas, and marine resources for sustainable development*. Available at: <https://unchronicle.un.org/article/goal-14-conserve-and-sustainably-use-oceans-seas-and-marine-resources-sustainable> (Accessed: 1 Apr 2020).

Dang, H. et al. (2019), 'Editorial: the responses of marine microorganisms, communities and ecofunctions to environmental gradients', *Frontiers in Microbiology*, 10:115, pp. 1-5.

Doerffer, R. OLCI Level 2 algorithm theoretical basis document: Ocean colour turbid water. Geesthacht, Germany: GKSS, 2010. 50.

ESA (n.d.) Ocean Quality and Science Flags. Available at: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/level-2/quality-and-science-flags-op> (Accessed: 01 Dec. 2019)

Godo, O. et al. (2012), 'Mesoscale eddies are oases for higher trophic marine life', *PLoS One*, 7(1), pp. 1-9.

Gregg, W. and Casey, N. (2004) 'Global and regional evaluation of the SeaWiFS chlorophyll data set', *Remote Sensing of Environment*, 93, pp. 463-479.

Groom, S. et al. (2019) 'Satellite ocean colour: current status and future perspective', *Frontiers in Marine Science*, 5, pp. 1-30.

Holliday, D. et al. (2011), 'Incorporation of larval fishes into a developing anti-cyclonic eddy of the Leeuwin Current off southwestern Australia', *Journal of Plankton Research*, 33(11), pp. 1696-1708.

Hosegood, P. et al. (2017), 'Nutrient pumping by submesoscale circulations in the Mauritanian upwellings system', *Progress in Oceanography*, 159, pp. 223-236.

Hsu, P. et al. (2019) 'Vertical structure and surface patterns of Green Island wakes induced by the Kuroshio', *Deep-Sea Research Part I*, 143, pp. 1-16.

Isoguchi, O. et al. (2009), 'Investigation of Kuroshio-induced cold-core eddy trains in the lee of the Izu Islands using high-resolution satellite images and numerical simulations', *Remote Sensing of Environment*, 113, pp. 1912-1925.

Jones, A. et al. (2014) 'Fine-scale hydrodynamics influence the spatio-temporal distribution of harbor porpoises at a coastal hotspot', *Progress in Oceanography*, pp. 1-20.

Liu, Z. et al. (2019) 'Tempo-spatial variations in the Kuroshio Current in the Tokara Strait based on long-term ferryboat ADCP data', *Journal of Geophysical Research: Oceans*, 124, pp. 6030-6049.

Loveday, B. and Evers-King, H. (2019), 'Python Universal Sentinel Downloader' (Version 2.1) [Source code].

Mackenzie, B. (2000), 'Turbulence, larval fish ecology and fisheries recruitment: a review of field studies', *Oceanological Acta*, 23(4), pp. 357-375.

Mahadevan, A. and Tandon, A. (2006), 'An analysis of mechanisms for submesoscale vertical motion at ocean fronts', *Ocean Modelling*, 14, pp. 241-256.

Mahadevan, A. (2015), 'The impact of submesoscale physics on primary productivity of plankton', *Annual Review of Marine Science*, 8, pp. 161-184.

MATLAB 9.5.0 / R2018b and Mapping Toolbox 4.7, and Image Processing Toolbox 10.3, The MathWorks, Inc., Natick, Massachusetts, United States.

Morel, A. et al. (2007) 'Examining the consistency of products derived from various ocean colour sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach', *Remote Sensing of Environment*, 111, pp. 69-88.

O'Reilly, J. et al. (1998), 'Ocean colour chlorophyll algorithms for SeaWiFS', *Journal of Geophysical Research*, 103, pp. 24937-24953.

Palacios, D. (2002), 'Factors influencing the island-mass effect of the Galapagos Archipelago', *Geophysical Research Letters*, 29(23), pp. 1-4.

Platt, T. et al. (2009), 'The phenology of phytoplankton blooms: ecosystem indicators from remote sensing', *Ecological Modelling*, 220, pp. 3057-3069.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Rothschild, B. and Osborn, T. (1988), 'Small-scale turbulence and plankton contact rates', *Journal of Plankton Research*, 10(3), pp. 465-474.

Scales, K. et al. (2014) 'Mesoscale fronts as foraging habitats: composite front mapping reveals oceanographic drivers of habitat use for a pelagic seabird', *Journal of the Royal Society Interface*, 11, pp. 1-8.

Schwarz, J. (2019), 'AVISO Data Extraction 20200205' (Version 1) [Source code].

SNAP – ESA Sentinel Application Platform 7.0.0 and SNAP Sentinel-3 Toolbox, European Space Agency, Paris, France.

Suthers, I. et al. (2004) 'Entrainment and advection in an island's tidal wake, as revealed by light attenuation, zooplankton and ichthyoplankton', *Limnology and Oceanography*, 49(1), pp. 283-296.

Wang, Y. et al. (2016), 'Warming and weakening trends of the Kuroshio during 1993-2013', *Geophysical Research Letters*, 43, pp. 9200-9207.

Zeng, X. et al. (2017), 'Variability of Changjiang Diluted Water revealed by a 45-year long-term ocean hindcast and self-organising maps analysis', *Continental Shelf Research*, 146, pp. 37-46.

Zhang, Y. et al. (2019), 'Submesoscale and Mesoscale Eddies in the Florida Straits: Observations from Satellite Ocean Color Measurements', *Geophysical Research Letters*, 46, pp. 13,262-13,270.

Zheng, Q. et al. (2008) 'Sub-mesoscale ocean vortex trains in the Luzon Strait', *Journal of Geophysical Research*, 113, pp. 1-12.

Zheng, Z. and Zheng, Q. (2014) 'Variability of island-induced ocean vortex trains, in the Kuroshio region southeast of Taiwan Island', *Continental Shelf Research*, 81, pp. 1-6.

Zhu, X. et al. (2017) 'Tidal currents and Kuroshio transport variations in the Tokara Strait estimated from ferryboat ADCP data', *Journal of Geophysical Research: Oceans*, 122, pp. 2120-2142.